NOAA Video Data Management System (VDMS): No Video, No Oscar®!

A. Fiolek
NOAA Central Library

D. Anderson
NOAA Central Library

D.W. Collins
NOAA National Oceanographic Data Center

M.L. Cumberpatch
NOAA Central Library

Edited by:
J. Beattie, IPT Team Leader
NOAA Central Library

Abstract - The NOAA Central Library (NCL) staff were integral to the formation of the Video Data Management System (VDMS) Plan for the NOAA Office of Ocean Exploration (OE). The VDMS is part of the larger comprehensive OE Data Management Project. This plan addresses both digital video and digital image formats. In support the requirements to document, store, and provide access to video and image information the NCL initiated the OE Digital Video Pilot Project in 2002. The Pilot Project presently provides online access to more than 1000 MiniDV and 500 DVCAM tapes, over 1000 DVDs, and access to more than 100 digital video highlights clips online collected by NOAA cruises in ocean exploration. Metadata guidelines (DV12 and DI12) were developed to aid data managers and librarians in the creation of MARC21 standard metadata records.

An inventory of the OE Video Collections is available online at: http://docs.lib.noaa.gov/OEDV/Inventory_DigitalVideoData2002-2004_01_19_2005.doc. Examples included in the Pilot Project are:

- ***Islands in the Stream 2002 Expedition: Exploring Underwater Oases*** **collection, including 100 MiniDVs of ca. 100 hours of digital video footage and 101 digital photos. NCL provides online access to video highlights, video clips, annotations to the original video footage, digital images with annotations, and cruise reports.**

- ***Oil and Gas Cold Seeps in the Gulf of Mexico, 1992-1995*** **collection, including 30 DVDs of ca. 60 hours of annotated digital video footage. NCL provides online access to 29 digital video clips and cruise reports.**

- ***Life on the Edge 2003: Exploring Deep Ocean Habitats*** **collection, including 52 MiniDVs of ca. 60 hours of annotated digital video footage on MiniDVs generated from 17 dives by the submersible Johnson Sea-Link II. NCL provides online access to video highlights, annotations to the original video footage, digital images, cruise reports, and educational lesson plans for Grade 5-12.**

- ***Investigating the Charleston Bump, August 2-16, 2003*** **collection, including 35 MiniDVs of ca. 40 hours of annotated digital video footage, and over 100 digital photos. NCL provides online access to video highlights, annotations to the original video footage, microscopic video clips, digital images, cruise reports, and educational lesson plans for Grade 5-12.**

- ***Life on the Edge 2004: Ft. Pierce, Florida to Cape Lookout, North Carolina, June 8 - June 22, 2004*** **collection, including 100 hours of Johnson Sea-Link I footage and 30 minute highlights on MiniDVs. NCL provides online access to video highlights, annotations to the original video footage, digital images, cruise reports, scientific data sets, and educational lesson plans for Grade 4-8.**

Once the OE Digital Video Pilot Project demonstrates its initial capability to acquire, document, manage, access and preserve digital video and image data, the initiative will be coordinated across all NOAA Line and Program Offices. Over the next five-years (2006-2010), there will be multiple access points to NOAA video data:

- **Continued access to multi-platform video images through the NOAA Libraries Online Catalog (NOAALINC) and the Online Computerized Library**

Center's (OCLC) WorldCat catalog. WorldCat is the world's largest and richest database of bibliographic information, linking the catalogs of over 53,548 libraries from 96 countries around the world.

- **Complete the development of a "NOAA Video Library Portal," providing integrated Web-enabled "one-stop shopping" access to essential video data and video management across the Agency. This approach will overcome the presently distributed, multi-formatted data and information scattered throughout NOAA.**
- **Provide video data via a gateway link from NOAA National Oceanographic Data Center's server where the OE central catalog will reside. NCL will provide Z39.50-compliant standard mapping for the OE catalog gateway and maintain a Web link for input from other sources to merge video into other products.**
- **Permanently archive important NOAA video from all line offices.**
- **Realign and address current video data management gaps in the present, fragmented structure.**

I. INTRODUCTION

The NOAA Office of Ocean Exploration (OE) and the NOAA National Marine Sanctuaries Program (NMSP), among other NOAA Line Office components, are routinely creating digital and analog videos that document program activities and data collection tasks (e.g., submersible dives, cruise highlights, and clips). As required by NOAA Administrative Orders NAO 15-217 and NAO 205-17, the NOAA Central Library[1] (NCL) and National Oceanographic Data Center[2] (NODC) are receiving an increasing number of video data collections from diverse NOAA components, including OE, NMSP, and the NOAA Coral Reef Information System (CoRIS).

Beginning in late 2002, the NCL began collaborating with OE Data Managers to develop and implement a data management plan for OE. The Integrated Product Team (IPT) was created to develop the plan, with a working group assigned to develop documentation, determine a work flow process and establish 'best practices' to support OE video data management requirements. The IPT recognized that it would also be beneficial if the requirements, documentation and system could serve as a model for the whole agency. In coordination with OE, NCL and the NODC Coastal Ocean Lab (COL) are developing the Video Data Management System (VDMS) for acquiring, accessioning, cataloging, maintaining and accessing digital video data from NOAA Offices. This document describes procedures to assure that digital video data files from these and other sources are managed consistently and effectively for the long term with minimal staff resource requirements.

II. MEDIA AND CONTENT MANAGEMENT

Video and image data present many challenges for Principal Investigators (PIs), data managers, and metadata librarians who work with the video and images after the conclusion of the project. The VDMS system accesses video data captured on MiniDV, DVCAM, and VHS tape media or on DVD. Using a video processing workstation, video data in native video formats is encoded (converted) to current industry and archival standard formats (e.g., MPEG-2 or MPEG-4) to facilitate online access and for long-term management.

Two types of video data are typically archived at NCL. Original tapes containing the entire sequence of video shot provide a relatively complete record of events during a cruise, dive or other activity. Highlights and clips typically contain 15 second to 15 minute excerpts of interesting or unusual features. Highlights are usually selected and created by the cruise PI and/or Data Manager. At present, NCL provides online access to highlights and clips video segments through a link in records of NOAALINC, the library online catalog located at: http://www.lib.noaa.gov/uhtbin/Webcat. A search for "digital video online" will list all catalog metadata records containing one or more links to digital video clips, as well as other related media and documents. Fig. 1 displays an example of NCL metadata record in MARC21 standards.

Original video tapes are stored in the NCL Video Archives in a climate controlled room and will be migrated to new media as necessary for continued archival maintenance. As additional resources become available, online access to the original tapes will be implemented.

III. ORGANIZING VIDEO DATA

The IPT VDMS working group developed a set of tools and products to assure harmonized and standard access to valuable data and information collected during OE-sponsored cruises. These tools and products include:

- VDMS requirements document: This document defines technical standards for the system, describes archival storage formats and conditions, and specifies online retrieval requirements.

[1] NOAA Central Library, SSMC-3 2nd Floor, 1315 East West Highway, Silver Spring, MD, 20910. Email may be sent to vdms@noaa.gov.

[2] National Oceanographic Data Center, SSMC-3 4th Floor, 1315 East West Highway, Silver Spring, MD, 20910. Email may be sent to NODC.DataOfficer@noaa.gov.

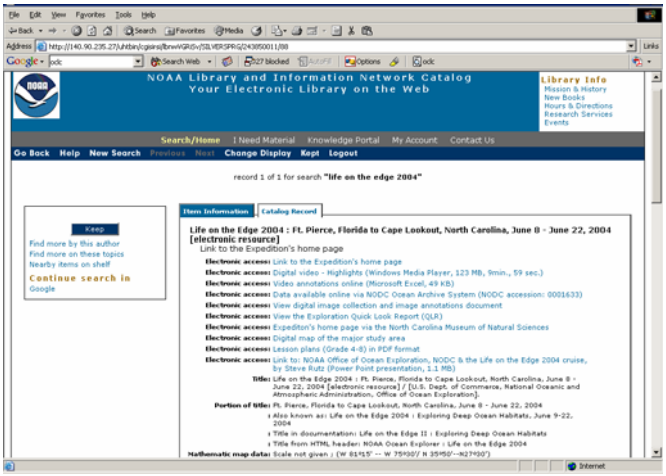

Fig. 1. Example of NOAALINC metadata record in MARC21 standards. Collection record describes Life on the Edge 2004 Expedition (full metadata available by searching NOAALINC at: http://www.lib.noaa.gov/uhtbin/webcat).

- Metadata standards requirements: The Federal Geographic Data Committee (FGDC) metadata standard supports descriptions of geospatial characteristics of the scientific data and MARC21 is the library-wide standard for documenting a resource (e.g., video or still image). Both the FGDC repository and NOAALINC databases used by the VDMS enable simultaneous data searches using the Z39.50 search protocol.
- Documentation guidelines: Each collection, tape, and highlights clip is being documented using guidelines provided by NODC/NCL. These guidelines are referred to as DV12 (Digital Video 12 descriptive elements) and DI12 (Digital Image 12 descriptive elements).
- DV12 and DI12 templates: These templates help Data Managers to create metadata records in both FGDC and MARC21 standards.
- Video data management workstation: NCL may provide a computer workstation to facilitate video editing functions for the collaborating investigators and data managers.

IV. DOCUMENTATION REQUIREMENTS

Imagine trying to understand a video created today in 50 years: When and where was the video taken, by whom, and for what reason? What is the format of the file and the encoding used to create the file? These few questions highlight the need for obtaining as much descriptive and technical metadata from the PI or data manager soon after the completion of each cruise. PIs and data managers provide the NODC/NCL metadata librarian with metadata for a video and/or image collection using the DV12 and/or DI12 templates, usually along with time-stamp oriented annotations for each individual video tape. They also typically provide copies of cruise reports and/or other data reports that were developed concurrent with the video collection. Descriptive information about the content of a video and the technical details about file formats, encoding algorithms, and processing equipment are needed to ensure that these videos are accessible and meaningful to future generations.

V. NODC/NCL PROCESSESS

Once video tapes and related data are received at the NCL, the metadata librarian creates a MARC21 record in the NOAALINC online database for the collection of video and related documents. Fig. 2 illustrates the workflow used by NCL for providing online access to NOAA video data [1].

NCL notifies NODC and, as described in [2], an NODC data content manager creates an accession entry in the NODC Accession Tracking Data Base (ATDB) for the collection of tapes, which assigns an NODC accession number as a tracking number for the collection.

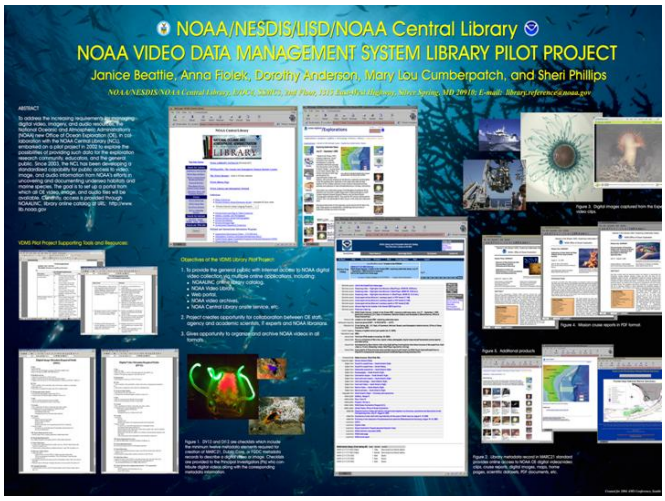

Fig. 2. VDMS poster outline of the NCL workflow [1] (full size image available online at http://docs.lib.noaa.gov/OEDV/VDMS_Poster.pdf).

Most ocean data archived at NODC can be discovered and downloaded using the NODC Ocean Archive System (OAS, online at http://www.nodc.noaa.gov/search/prod/). FGDC metadata are automatically harvested from the NODC ATDB accession entry for inclusion in the NODC Metadata Manager and Repository (NMMR) database. Other, non-video data collected during the cruise with video data may also be placed in the same NODC accession or a separate accession, with a reference to the video accession. Fig. 3 illustrates the generic flow of video and other data from data collectors and originators to the archive centers. The end-to-end process is more fully discussed in [3].

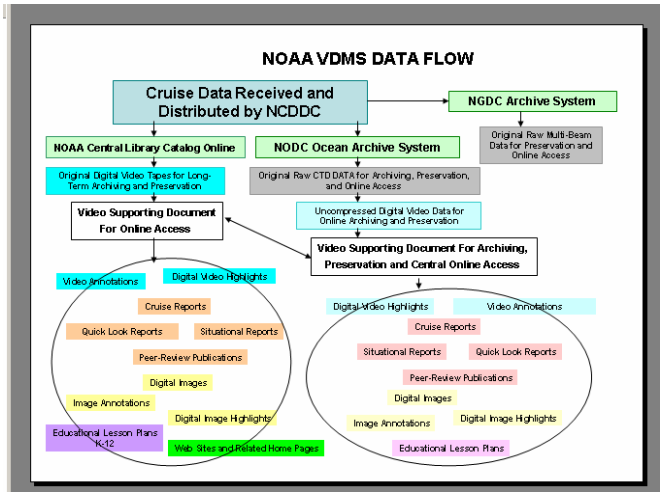

Fig. 3. Schematic diagram showing generic data flow in the NOAA VDMS (full size image available online at http://docs.lib.noaa.gov/OEDV/VDMS_Workflow_Ocean2005.jpg).

VI. FUTURE ACTIVITIES AND SUMMARY

By leading the VDMS project and being a major participant in the NOAA Ocean Exploration data management strategy, the NOAA Central Library team continue to collaborate closely with NOAA scientists, oceanographers, and IT specialists working within the OE Project to develop data management requirements and strategies. This project provides an ongoing opportunity to improve the quality and completeness of metadata and information used in the NOAALINC catalog and to provide online access to NOAA ocean exploration video and related data to a global customer base.

Future activities include improving in-house video processing capabilities -- including encoding raw video data into online-accessible versions, setting up an informal seminar series that highlights video collections, and examining how other groups (e.g., educators, other scientists) are using digital video data from NOAA.

The VDMS project is a good foundation of procedure for NOAA offices to archive and preserve the agency's scientific video data in both physical and online formats for future generations. As NOAA video becomes more widely available through the VDMS, perhaps it is time to establish a NOAA "Oscar®" for "the best online scientific video!

ABBREVATIONS

MARC21 (MAchine Readable Catalog) – Standards for the representation and communication of bibliographic and related information in machine-readable form.

OCLC – Online Computer Library Center, Inc.

REFERENCES CITED

[1] D. Anderson, J. Beattie, M.L. Cumberpatch, A. Fiolek, Digital video needs for oceanographic images for the National Oceanic and Atmospheric Administration (NOAA): Phase 2. *Proc. Oceans 2003*, Paper #148, pp. 1292-1296, 2003.

[2] J. Beattie, A. Fiolek, D. Anderson, M.L. Cumberpatch, S. Phillips, NOAA Video Data Management System, Library Pilot Project. 2004. Poster session at 84th American Meteorological Association conference, January 2004. Accessible online at http://docs.lib.noaa.gov/OEDV/VDMS_Poster.pdf (last accessed on 6/30/05)

[3 D.W. Collins and S.B. Rutz, The NODC archive management system: archiving marine data for ocean exploration and beyond. *Proc. Oceans 2005*, Abstract #050520-02, in press.

[4] S. Mesick, E. Roby, S. Gottfried. End-to-End Data Management: a success story. *Proc. Oceans 2005*, Abstract #050408-05, in press.